

# UNCLASSIFIED

AD NUMBER
AD434732
NEW LIMITATION CHANGE
TO Approved for public release, distribution unlimited
FROM Distribution authorized to U.S. Gov't. agencies and their contractors; Administrative/Operational Use; MAR 1964. Other requests shall be referred to Office of Naval Research, Arlington, VA 22203.
AUTHORITY
ONR ltr dtd 2 Sep 1970

THIS PAGE IS UNCLASSIFIED

**UNCLASSIFIED**

**AD**

**434732**

**DEFENSE DOCUMENTATION CENTER**

**FOR**

**SCIENTIFIC AND TECHNICAL INFORMATION**

**CAMERON STATION, ALEXANDRIA, VIRGINIA**



**UNCLASSIFIED**

NOTICE: When government or other drawings, specifications or other data are used for any purpose other than in connection with a definitely related government procurement operation, the U. S. Government thereby incurs no responsibility, nor any obligation whatsoever; and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data is not to be regarded by implication or otherwise as in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use or sell any patented invention that may in any way be related thereto.

434732

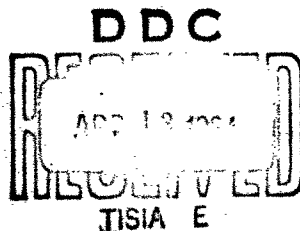
CATALOGED BY DDC  
AS AD NO. \_\_\_\_\_

THEMATIC EVALUATION OF MANAGEMENT POTENTIAL

FINAL REPORT

Prepared for:

Office of Naval Research  
Department of the Navy



By:

Personnel Research and Development Corporation  
Cleveland, Ohio

March, 1964

NO OTS

### LIST OF TABLES

<u>Table</u>	<u>Title</u>	<u>Page</u>
1	Order of Scales in Original Manual	11(a)
2	Time taken to Complete Evaluations	22(a)
3	Basic Reliability Indices for Common and Uncommon Cases	25(a)
4	Percent Disagreement by Scale, 125 Uncommon Cases - Data Calculated from Scatter Plots	26(a)
5	Distribution of Ratings by Retained Scale for the Five Cases in the Common Population	26(b)
6	Relative Frequency and Standard Score Values of the Retained Scales (as revised)	28(a)
7	Intraclass Correlation by Rater Type Common Cases - 5 Cases Rated by all Raters	28(b)
8	Intraclass Correlation by Rater Type Uncommon Cases - Each Case Rated by Two Raters	29(a)
9	Intraclass Correlation Between Rating Scales	29(b)
10	Original and Rotated Orthogonal Factors of 18 Rating Scales in TEMP	30(a)
11	Order and Reliability of Scales in Revised Manual	31(a)

THEMATIC EVALUATION OF MANAGEMENT POTENTIAL

Final Report

Erwin K. Taylor  
Edwin C. Nevis  
Richard W. Wallen

This Study was Supported in Part by  
Office of Naval Research Contract Nonr-4016(00)  
And  
U. S. Public Health Service Grant M-2158

Reproduction in whole or in part  
is permitted for any purpose of the  
United States Government

March, 1964

Personnel Research and Development Corporation  
Cleveland, Ohio 44120

Copyright 1964

Personnel Research and Development Corporation  
Cleveland, Ohio

Reproduced in the United States of America

## TABLE OF CONTENTS

	Page.
I Perspective of the Research . . . . .	1
II TEMP - Its Development and Method . . . . .	6
III The Research Design . . . . .	18
IV Results . . . . .	25
V Revision of the Manual. . . . .	30
VI Uses of the Manual. . . . .	31
VII Further Research. . . . .	35
VIII Summary and Conclusions . . . . .	37
Appendix A	
Appendix B	
Appendix C	
Bibliography	



## I. PERSPECTIVE OF THE RESEARCH

### A. The Importance of Personality Variables in Managerial Assessment

The need for sensitive measurement of normal personality differences is quite apparent to those who have devoted a good deal of attention to the problems of psychological assessment for the purposes of predicting job effectiveness. To date, our ability to measure aptitudes and proficiencies is considerably better than what we have been able to accomplish in the field of personality assessment. However, daily experience and research findings indicate that problems of job adjustment and job failure stem more from personality problems than for want of intelligence or technical skills. A survey by F. J. Gaudet (10) indicates that, among management and executive level personnel, seven times as many failures are the result of personality problems than are due to a lack of technical skills necessary to handle the job.

To meet this need for effective screening and appropriate matching of personality characteristics and job requirements, many techniques and approaches have been utilized. The ordinary questionnaire approach, while still frequently used, has not demonstrated itself to be very useful in this regard. This is particularly true with executive level personnel, or where motivational factors may lead the applicant to distort his replies. Hope of improving this situation by the development of forced-choice, self-description forms, such as those of Edwards, Gordon and Ghiselli, is waning, and the need for more penetrating and yet economical methods of evaluation continues. This study is an attempt to modify the use of the Thematic Apperception Test (TAT) to meet this need in the assessment of managerial potential.

### B. The Use of Projective Techniques in Managerial Assessment

The use of projective techniques in managerial assessment rests squarely on the assumption that personality factors are of utmost importance in managerial performance, and that a manager's style of performance is reflected in every assignment he undertakes. The higher the management level, the greater the ambiguity and lack of clarity in the problems presented, making for considerable latitude on the part of managers in decision-making. Managerial behavior is based to a large extent on typical work attitudes and habits, general outlook toward life and work, value systems, and other motivational considerations.

Higher level jobs are composed of a complex, interwoven variety of interpersonal relationships. It is difficult to define personality requirements for these jobs in specific terms, such as in the case of technical skills. Much of management work requires that relatively unstructured situations be clarified and that order be developed out of complex, unclear stimuli. Unstructured measuring devices, such as projective techniques, provide a highly useful way to obtain a sample of a manager's work style and of his orientation to managerial relationships, rewards, and frustrations.

Projective techniques, such as the TAT, have been routinely used in psychological clinics, but have not been widely used in the evaluation of industrial personnel. One reason is that, in their original form, most projective methods are too lengthy for convenient use as personnel tools. Furthermore, a professionally trained psychologist must administer and interpret the productions of the assessees. Psychometric tests, on the other hand, can be administered and scored by non-professionals. Thus, the substantially greater cost of projective testing has hindered its widespread use. A third disadvantage of projective tests is the fact that they do not appear to have obvious relevance to business activities. To a salesman who knows nothing of psychology, the administration of a test with the title "Sales Aptitude" is quite understandable. But he often has trouble understanding why he should draw pictures or interpret ink-blots. This lack of "face validity" and the public's belief that these tests are for "sick people" make some industrial psychologists reluctant to use projective methods in personnel assessment.

Another barrier to the widespread industrial use of projective techniques is the rather considerable difficulty encountered in attempting to validate them. Unlike psychometric tests, the information they yield has not been couched in terms amenable to the usually accepted techniques for statistical analysis. It is possible, of course, to impose objective scoring schemes on the rich and varied responses, but these schemes have turned out to be tedious, time-consuming, and often not very relevant to the validation problem. The usefulness of projective techniques depends greatly upon the training and experience of the interpreter. It is the validity of the interpretation that is in question rather than the validity of the subjects' responses. In projective methods, then, the interpreter becomes an inherent part of the technique and, as in the case of the

interview, we must necessarily become concerned with the validity of the individual interpreter. This creates a difficult situation for personnel psychologists reared in the American tradition of objective scoring methods and clearly defined cutting scores.

Despite these barriers, however, projective techniques are gradually finding a place in the testing of industrial personnel. Experienced psychologists are finding that projective methods permit them to make significant inferences about personality and performance that are not possible with self-descriptive questionnaires. This study is an attempt to overcome some of the barriers and to put the use of projective techniques in personnel assessment on a more sound footing.

#### C. The Development of the TAT as a Personnel Assessment Device

In 1935, Morgan and Murray announced the first version of the TAT (22). It was used as one instrument in an extensive personality study conducted by Murray and his co-workers; and in 1938, this work resulted in the publication of Explorations in Personality (23).

A number of psychologists soon became interested in the method and, with the advent of World War II, it was tried out in a number of military applications, both clinical and non-clinical. The original test materials consisted of 30 pictures. From these, four sets of 20 pictures could be constructed, one each of girls, women, boys, and men. The subject was asked to make up a story about each one of these pictures. Over the years, there have been substitutions and modifications in the original set, but the pictures now in use probably will remain for some time to come. Originally, it was customary to administer the full 20 card set, using two testing periods, each about an hour long. The pictures presented during the first test period are numbered from 1 to 10 and are, in the main, somewhat clearer and less vague and fantastic than the 10 pictures presented during the second testing session. Through the years, practical demands on the time of psychologists have resulted in shortening the number of cards used so that, today, few psychologists use the full 20 cards.

After the War, full and complete manuals presenting interpretive principles, along with examples, began to appear. Tomkins (31) published his approach to the technique of interpretation in 1947. In 1949, another complete manual was published by Betty Aron (2). A manual by Stein (28) appeared in 1948 and was

revised in 1954. In 1956, Henry (17) published the Analysis of Fantasy, which not only developed his interpretive principles, but gave extensive examples of records produced by people in other cultures.

While these manuals were appearing, a substantial amount of research work was being published in the psychological journals. The TAT was being subjected to validity tests. Special groups, such as schizophrenics, crippled children, stutterers, and ulcer patients, were studied, and their characteristic responses to the TAT cards described. Gradually, the common responses to cards became known, and a considerable body of information had accumulated to aid the beginning projective analyst.

Not all research with this technique has proven successful. The efforts of the U. S. Army to devise a force-choice military version of the TAT did not demonstrate predictive validity. The unpublished efforts of the personnel research department of a large industrial organization to construct an "industrial TAT" also failed to come to fruition. The problem in the latter case appears to have been that of selecting overly structured pictures, yielding too many descriptions of the cards rather than "projective" protocols.

In 1948 and 1949, Gardner (9) and Henry (15) published the results of their studies of business executives. They had found that certain characteristics, some of which could be inferred from the TAT, differentiated successful from unsuccessful executives. Since then, relatively little has appeared in print on the use of the TAT in industrial situations. But some psychological consultants have been using it as a regular part of an assessment test battery.

The present investigators have been using the TAT in personnel assessment work for almost 15 years. This use has been in the form of "blind analysis," in which the technique is but one part of an assessment battery including objective tests and interviews. While the TAT does not have face validity as such, it has been found to be more acceptable to industrial personnel than the Rorschach, "Draw-a-Man," and most other projective devices. It is often seen as a challenge and hardly ever seems to engender embarrassment or feelings of "childishness." It yields much richer material than do personality inventories where higher level personnel are to be evaluated. The extent to which analyses have coincided with independent observations of professionally trained interviewers and with observations of job performance has been encouraging. Current studies of the validity

of the entire assessment procedure likewise appear very promising. Thus, sufficient practical experience in using the test with managerial people has accumulated to permit the formulation of the interpretive technique developed in this study.

#### D. Orientation of This Study

This modification is shorter than the original Thematic Apperception Test (TAT), and it uses written stories rather than spoken ones. A psychometrician can administer the test, thus releasing the time of professionals to do interpretation. These changes, along with several others, require a new approach to the interpretive problem, although most of the present knowledge about TAT interpretation is applicable. We call this adaptation the Thematic Evaluation of Managerial Potential (TEMP).

The study grew out of the approach used and the data obtained in psychological assessments by the Personnel Research and Development Corporation for industrial organizations over the past nine years. In such assessments, it has been standard practice to secure written protocols on 10 TAT cards. The resulting stories, together with a sentence completion test and certain personal history information, are supplied to a psychologist who, using only these data, prepares a narrative report concerning the individual's work attitudes and habits, drive and ambition, intellectual effectiveness, emotional resources, interpersonal relations, and overall qualifications for the purpose for which he is evaluated, be it selection, promotion, or counseling and development reasons. In addition to the projective portion of the assessment, a battery of paper and pencil aptitude and personality tests, a comprehensive personal history form, and independently conducted interviews are utilized. Clients typically employ these assessments at the key personnel levels, including salesmen, engineers, accountants, foremen, department heads, and top level executives.

As is typical of most assessments, in either the clinical or industrial setting, analysis of the projective materials in this situation is based generally upon a free-ranging, rather intuitive approach in which the analyst reviews the material, responds implicitly to certain cues, begins to put the pieces together and to formulate his hunches into an overall description and set of predictions. The task is viewed primarily as one of formulating the personality dynamics of the

individual, carrying with it a somewhat clinical orientation, and turning this into a portrait in words which describes the assessee's potential behavior in the job and work setting for which he is being studied.

In order to make more direct predictions concerning relevant job behavior patterns and to learn more about the value or validity of the TAT in this type of assessment, the present approach places its emphasis upon the determination of a number of specific behavior rating scales by means of explicitly stated indicators which presumably are predictive of the particular set of job behaviors in each rating area. The goal is to obtain a method which combines the clinician's rich, integrative approach to personality with the values of quantification. The point of departure is the job behavior to be rated -- not a system of personality classification or an a priori scoring scheme. The evaluation system is built around a framework of behavior variables considered important to predictions about the normal person and his adjustment to the world of work. The manual for this system was constructed to be a guide for the rating of 18 variables (or scales) concerning various aspects of managerial performance. The aim was to start from consideration of what is deemed necessary to understand and predict in the work setting, and then to develop a means by which such evaluation can be accomplished. In this way, it was hoped to make the system as practical as possible and to get away from the traditional attempts to adapt clinical concepts to the evaluation of essentially well-adjusted individuals.

This report presents the system (manual) developed and the results of a study to meet two major aims with this system. These aims are: (1) to determine if written protocols to selected TAT cards can yield meaningful discriminations along a variety of behavior variables for selected normal subjects, and (2) to provide a standardized interpretation procedure by means of which these variables may be evaluated -- that is, to obtain a semi-objective procedure which would yield high reliability among adequately trained observers.

## II. TEMP - ITS DEVELOPMENT AND METHOD

### A. The Present Method of Administration of the TAT

Purpose - The purpose of TEMP is the same as that of the TAT: To learn about the important aspects of the subject's personality. Since the subject does not know exactly what the tests analyst is looking for, it is difficult for him

to suppress or distort his responses in any systematic way. Of course, he can suppress content that he thinks is unacceptable, e.g., shocking sexual or sadistic content and, in fact, we expect to find such suppression as a resource in a well-integrated person. But, he cannot know that the "tired boy" he sees in a picture is more often seen as a grief-stricken man. Nor does he realize the marked difference between a story that ends with a boy hoping to leave home and one that ends with the boy actually leaving.

This ability to by-pass the conscious intention and control of the subject is one of the valuable features of the test. While stories can be changed somewhat by subjects, they are less susceptible to intentional control than are self-descriptive questionnaires for studying personality. Thus, in an employment or assessment situation where the desire to make a favorable impression is strong, the projective test is particularly useful.

Administration - In TEMP only 10 of the TAT cards are used. These are:

Card 1	Card 8BM
Card 2	Card 9BM
Card 3BM	Card 12M
Card 6BM	Card 17BM
Card 7BM	Card 20

These 10 cards were selected on the basis of years of experience in use of the TAT in industrial assessment, showing that these cards generally yield the most useful material for work-oriented predictions. The 18 variables to be rated appear to be better evaluated by these cards than by means of any of the other cards in the original TAT set.

The materials needed for administration of TEMP are as follows:

1. Ten TAT cards mounted in plastic.
2. Ten sheets of plain, unlined paper stapled together.
3. A pencil with a good eraser, preferably a Number 2 or medium lead.
4. A direction sheet.

The subject is seated in a quiet, well-lighted room at a table or desk, in a comfortable chair. The 10 cards are placed in front of him on the desk, arranged in numerical order. Card 1 is on top.

The examiner says:

"This is a test of your imagination. You have been given a set of 10 cards, each with a picture on it. Your task is to write a story about each picture. In each story, we would like you to tell what is happening in the event pictured, what led up to it, what the outcome will be, and what the characters are thinking and feeling. Are there any questions?"

If not, the examiner proceeds:

"This test is not timed. Write the stories, one story per sheet of paper, on one side only. Each story should not take longer than about ten minutes to write. There is a number on the back of each card. (Examiner turns the top card over and points out the location of the number on the back of the card to the subject.) Please put that number on the sheet you use for that story. Be sure to keep the cards in the same order in which they were given to you. I will return to check your first story in about ten or fifteen minutes."

No further stories, after the first one, are checked by the examiner.

The examiner should leave a direction sheet for the subject to use as a reference. This sheet reads as follows:

This test consists of a series of cards on each of which there is a picture. Your job is to write a story about each picture. In each story, please tell what is happening now, what led up to the present, and what the outcome will be. Also include what the characters are thinking or feeling.

Each story should be no longer than one written page and should not take longer than about ten minutes to write. Please follow the order of the cards as presented.

The examiner will check with you after you have completed the first story to see if you have any questions. Do not hesitate to ask any questions you may have now or after you have completed the first story.

After allowing ten to fifteen minutes for completion of the first story, the examiner checks to see that the four requirements have been included and that the story follows a logical procession. If the story is too long, the subject is asked to spend a little less time on the rest of his stories. If



the story is too short, he is asked to elaborate a little more and to make sure that all requirements are fulfilled. It is sometimes necessary to repeat the four requirements. The subject is not asked to re-do his first story, only to do the things suggested on the following stories.

This procedure differs from the standard TAT administration, and it is worthwhile to make some comparisons between the two. In giving the TAT, the examiner is present, listening to and recording the stories. In the TEMP, there is minimal contact with the examiner, and the stories are written by the subject. The examiner cannot inquire into unclear remarks or encourage compliance with the instructions after the first card. On the other hand, the relative exclusion of the examiner's personality makes conditions of administration more uniform for all subjects than in typical clinical uses of the TAT. In the TEMP, a rough limit is placed on the length of the story, while in the TAT the subject theoretically may continue as long as he wishes. This limitation sometimes results in less elaborate stories. At the same time, however, it forces the writer to plan and organize more than is required by oral stories.

To people who feel inadequate in spelling and grammar, writing poses a greater stress than speaking. However, when writing stories, the subject has more time to consider the organization of his stories and the "fit" of his sentences than when speaking them. Consequently, we expect somewhat better organization and less rambling in written stories. Traces of indecision or hasty impatience remain in the form of erasures and cross-outs. Then too, weaknesses in spelling, grammar, and punctuation can become painfully evident.

Using ten cards instead of the original 20 limits the range of situations that can be explored, but it increases the practicality of the test for industrial purposes. Most men can finish the test in about an hour and three-quarters. Less than an hour and one-half is considered "rapid" time, and more than two hours is regarded as "slow." The amount of material yielded in this time is quite adequate for interpretation; at least this is true for subjects who are of management caliber.

One of the major advantages of the written approach is to make possible economical administration. The examiner is needed only at the beginning of the two-hour period, and even this phase may be handled by a skilled psychometrist.

## B. Development of the Behavior Variables

The initial research leading to the development of standard procedures for evaluating written protocols to the TAT was undertaken with the aid of a U. S. Public Health Grant made to the senior author while he was on the faculty at Western Reserve University. The objective of this phase of the research was first to identify a number of variables believed by competent authorities to be pertinent to the evaluation of performance and potential of normal individuals in the business situation. The second objective was to determine the kinds of cues that competent and experienced TAT analysts would use as a basis for evaluating the finally selected variables.

Initially, a rather sizeable group of project consultants were employed for the purpose of providing tentative variables from a variety of sources. This committee of consultants submitted in excess of 1,000 suggestions for behavior characteristics that might be useful in evaluating individuals for managerial and executive level positions in business and industry. There was, of course, considerable overlap. At this point, a smaller group of four psychologists and a psychiatrist with considerable psychological training was constituted. \* This committee reduced the variables to 26 in number, and the project staff agreed to these as a workable beginning for the development of a system.

The committee of consultants and the project staff did some trial evaluation to see whether the 26 variables could be measured, using the traditional, global approach. This led to further screening of the variables and reduction of the final number to 18, which were unanimously accepted to be amenable to evaluation from written protocols. A larger number of variables might readily have been selected, but both the project personnel and the consultants agreed that to go beyond this number would probably involve exceeding the capacity of evaluators to make discriminations.

The 18 selected variables provided the basis for the construction of a

(\*) The consultants on this committee were Dr. William E. Henry and Dr. Donald Fiske of the University of Chicago, Dr. Alvin Scodel of Ohio State University, Dr. Boyd McCandless of Indiana University, and Dr. Jay B. Cohn. The first four are psychologists by profession and the fourth a psychiatrist

similar number of rating scales varying in length from four to six steps. A list of the titles of the 18 rating scales is provided in Table 1. Appendix A presents the complete scales for all the variables.

Before beginning the development of the specific indicators, a definition of each variable was prepared. This set forth the theoretical basis and something of the personality dynamics inherent in the behavioral steps of the scales. This descriptive material, useful in "anchoring" the scale steps, was included in the manual as an introductory section.

Having made the commitment that these 18 variables could be evaluated on the basis of written protocols from 10 TAT cards, the next step was to commit the basis for such evaluations to paper in the form of cues or indicators for evaluating each step of each scale. This was accomplished by the project staff, whose work was reviewed by the committee of consultants. The latter made a significant contribution in their critical review of the staff's work.

### C. Developing the Indicators

#### Interpreting the TAT

Ordinarily, the TAT is used to provide a fairly comprehensive description of a personality. It can yield data bearing on the major motives of the narrator, as well as his conflicts, his anxieties, and his techniques for reducing these anxieties and conflicts. Many of the inferences that are arrived at concern characteristics that are partially or totally unconscious and are, by definition, unknown to the narrator.

To make a comprehensive interpretation of a TAT record requires familiarity with some kind of personality theory, as well as familiarity with the ways in which various personality characteristics manifest themselves in TAT stories. Considerable experience and training are required, therefore, for the usual method of interpretation.

A recurring problem in work with the TAT is how to decide which of the various trends shown in the stories will actually appear in overt behavior. For example, suppose that the stories show a number of instances where a male character acts in a hostile manner toward a female character. Does this permit us to predict that the writer will treat female secretaries in the same way? While there is a great deal to be learned about answering such questions, it seems at present that the answer depends on a number of other characteristics

TABLE 1

Order of Scales in Original Manual

1. Accuracy of Perceptual Interpretation
2. Amount of Productivity
3. Consistency of Productivity
4. Whole vs. Detail Approach to Problems
5. Originality of Expression
6. Emotional Control
7. Involvement in Work
8. Need for Clear Assignment and Direction
9. Need for Emotional Support in Work
10. Reaction to Frustration and Failure
11. Level of Ambition
12. Intensity of Ambition
13. Relations with Superiors
14. Ability to Organize and Direct Others
15. Consideration for Subordinates
16. Insight into Others
17. Social Skills
18. Social Dominance

of the record. In the example used, if there were signs of impulsive behavior and of irritability, there is a greater likelihood that the writer would openly display hostility toward secretaries. The usual method of interpretation, then, must depend upon many cues or indicators that are not directly related to the behavior we wish to predict. It is not possible to make the simple assumption that the behavior of the central character represents directly the behavior to be expected from the writer of the stories.

In this study, we attempted to develop a different approach to interpretation, one that can be used by people who do not have the requisite training for the usual kind of interpretation. In order to advance beyond the global, intuitive, and artistic approach to interpretation, we have attempted to make public and explicit the various cues which would indicate a particular rating on some specific variable. Some difficult problems are encountered in reaching this aim, and we certainly cannot claim that our list of indicators is definitive. Our approach to this problem will be discussed later.

#### Kinds of Indicators Used

The most obvious, but perhaps not the best, indicators come from a content analysis of the stories. This includes the general mood, as well as the kind of outcome, that the narrative conceives. What are the characters in the stories striving for? What kinds of relationships are depicted? What barriers or constraints are presented to the central character? Of particular interest are contrasting stories. For example, in one story the central character may be successful in reaching some goal, while in another he may be unsuccessful. How can we account for the difference between these two outcomes? Does the writer appear to ascribe it to the character structure of the central figure, to the kind of goal he has set, to his available emotional supports, or to interference by the social or physical environment? Another example of contrasting stories is a case where one story depicts a friendly and supportive relationship between an older and a younger person, while another story depicts a relationship full of hostility and suspicion. Such contrasts can reveal the writer's assumption about various kinds of interpersonal relationships.

Another aspect of interpreting story content is the relationship between the generally accepted stimulus properties of the cards and the kind of story told by a particular writer. Discrepancies between the nature of the stimulus

picture and the written story may show anxiety, impulsiveness, reaction formation, or other personal characteristics of the writer. Card 17BM (man on rope), for example, ordinarily stimulates writers to think about contests, gymnastic exhibitions, or circus performers. If a writer tells a story about a man climbing down a rope to get off a water tower he has just painted, we are entitled to suppose that something about the stimulus picture disturbed the writer. One plausible guess is that he dislikes competitive or exhibitionistic situations. Of course, other features of the record must be examined before accepting this as the most probable hypothesis. It could equally well turn out that the writer's disturbance was occasioned by his distaste for and rejection of displays of physical strength.

A second kind of indicator comes from the form and style of the story. This includes such characteristics as the length of the story, the clarity of its organization, the kind of language that is used, the elaborateness with which details are added, and whether suspense is employed. We must remember that the writer is communicating with someone unknown to him and expects that his stories will be read. Particularly in situations where the test is used as a part of a selection battery, the writer would be keenly aware of the need to make a good impression. Therefore, it can be assumed that he would be doing the best job of communication and imagination of which he is capable. At the same time, he is less likely to control the style and format of his stories than he is to control the kinds of characters introduced and the outcome. For these reasons, indicators from the form and style of the stories are of great importance. Consider, for example, a record in which all stories start with a paragraph describing the picture. "This picture shows a boy, a violin, and something that looks like a sheet of paper. Presumably the boy is seated. He appears to be looking at the violin. His expression is thoughtful." This approach to stories is a direct sample of the writer's behavior. We can see that he is approaching this task cautiously, making a careful survey of the problem, and hesitates to begin the story. Furthermore, he somehow feels that he must share his observations with the eventual reader of the stories. This manner of approach is a far more powerful indicator of a cautious, detail-oriented person than when central characters are described in such a way as to indicate that they are cautious and detail-oriented. What we are saying is essentially that it is profitable to look at the stories as a work sample of written interpersonal communication.

A third source of indicators in written records is the grammar, spelling, and punctuation of the stories. These matters obviously have some relationship to educational level and to interest in writing. Spelling errors are common, however, even among college educated people. When these errors occur on simple words, they often indicate some tension or anxiousness. An example would be spelling "here" for "hear." Some jobs require the incumbent to write well, and in these cases the written record provides a sample of the writer's ability in this respect. Punctuation is usually sparse and often erroneous, even in the stories written by college graduates. A common error is the omission of quotation marks when writing dialogue. Careful attention to punctuation increases the probability that the writer is painstaking and detail-oriented.

A fourth set of indicators is derived from characteristics of the act of writing. These include such things as erasures, strike-outs, retracing of letters, and the omission of words or endings of words. These may reveal tension, impulsiveness, excessive meticulousness, or lack of attention to detail. Sometimes such inefficiencies are found almost entirely in one story. That would suggest that some stimulus characteristic of this card was particularly disturbing to the writer.

Another source of information about the writer comes from contrasting stories near the end of the record with stories at or near the beginning. For men who are accustomed to dictating reports and correspondence, and for those who have been out of school for a number of years, the writing of these stories is a formidable and fatiguing task. The effects of writing fatigue, as well as exposure to a number of emotionally-toned pictures, are usually a lowering of self-discipline and inhibitory controls. Consequently, the later stories in a record often differ from the earlier stories. The later stories tend to be somewhat briefer, somewhat less well-organized, and a bit more spontaneous. If anxiety is mounting, the characteristic "security operations" of the writer may appear in a very pronounced form in the later stories. For example, men who are cautious and rather obsessional may show no indecision on the first few stories, but later stories will show much qualification and indecision as to the age, feelings, occupation, or aims of the characters.

### Techniques for Developing the Indicators

In developing the indicators for TEMP, members of the project staff used both empirical and theoretical approaches. The staff members had considerable previous experience with projective techniques in a variety of clinical, counseling, and selection situations. This background gave them some ready-made ideas and hunches about the ways in which various personal characteristics were revealed in the written records. In addition, specific material for this project was obtained through a laborious analysis of cases from our own files. The available data had been collected from management and sales people who had been assessed over a period of years. Besides the written TAT record, a number of other kinds of information were available: reports written by interviewers, a personal history form, an incomplete sentence test, and a variety of tests of cognitive functioning. These voluminous data enabled the staff to develop and test hunches in an extremely flexible way.

One approach to developing indicators was largely empirical. For example, it would be noted that records of salesmen often contained quoted material -- that is, dialogue or monologue was used. This, in turn, suggested that people who were more at home in face-to-face situations would use quoted material. Since this was related to social skills in face-to-face situations, the use of direct quotations in the stories became an indicator for one level of the rating scale for social skills. A similar attack was made on the problem of why some writers gave names to the characters and others did not. In this case, the staff speculated that the use of names probably indicated an interest in people as individuals. As we searched through our records, it became clear that this formulation was too simple. For example, we found instances where characters had been given humorous but demeaning names. This seemed to indicate a contempt for people -- at least for those of low status. In a few instances, names were used but seemed to be inappropriate to the age or social class of the character. Sometimes, the names used were both appropriate and somewhat unusual, suggesting that the writer had made a strong effort to be creative and search for novelty. Thus, as with most other indicators, subtle differences in the handling of the indicator may make substantial differences in interpretation. Nevertheless, we decided to include the use of names for the characters as an indicator of social skills.



Another approach to selecting indicators is one that involves the use of psychological theory. In this approach, an attempt is made to relate certain personality patterns to the various scales. We assumed that people with an obsessive-compulsive character structure would be hard workers, motivated by a sense of duty. We also assumed that they would tend to suppress spontaneous emotional expressions. Thus, the indicators for an obsessional personality structure could be used on both the scale for involvement with work and the scale for emotional control.

A specific example of this kind of thinking can be taken from the indicators of the scale "Insight into Others." For Rating "E" indicators of distorted overconcern, we assume that people with paranoid trends would show this kind of reaction. One of the most potent indicators of paranoid trends is a kind of quibbling concern about the exact expression on the faces of the characters. Consequently, this was included in the list of indicators for Rating "E." At the same time, we assumed that the likelihood of distorted over-concern would be greater if the individual with paranoid trends were tense and anxious. One fairly specific indicator for tension is erasures, retracings, and cross-outs. That indicator was also included. There are, of course, people whose distorted over-concern is phobically-based rather than paranoid. Such people are likely to repress their aggression and hostility. Since such people do not have their aggression available for self-defense, they are often quite concerned about maintaining the good will of those around them. The amount of kindness or cruelty which they see in others is an important determiner of their sense of security. This led to the last indicator for Rating "E:" "Relationships among characters are often described in terms of how much or how little they like one another, or frequent comments about whether a character is 'kind-looking' or 'mean-looking'."

Since more than one personality structure may lead to the behavior which is designated by one rating on any of the scales, some of the indicators seem to be contradictory and inconsistent. This apparent inconsistency is due to our attempt to include indicators from more than one personality configuration. The reader who is familiar with types of personality, such as cycloid, schizoid, hysteroid, obsessional, and paranoid, will be able to see the reasons for the choice of many of the scale indicators.

The indicators as they stand, then, are mixtures of empirically-derived and theoretically-derived ideas. We have certainly not exhausted the possibilities for additional indicators. Professional psychologists may well wish to add some of their own discoveries to our list.

#### D. The Format of the Manual

The purpose of developing cues for evaluating the written protocols was, of course, to enable the project staff to design a manual which would enable personnel people to evaluate a set of protocols without the complete technical training now required of TAT analysts. Two obvious approaches to the manual format presented themselves. The manual could have been constructed on a card-by-card basis in which the indicators of style, content, and form for each of the 10 cards could be separately given on a scale-by-scale basis. This, we felt, would have led to excessive paging back and forth in the manual as its user endeavored to evaluate a protocol. It would also have led to an atomistic approach rather than a wholistic approach, making for discontinuity in arriving at each of the evaluations. Therefore, the second alternative of considering the scales one at a time was taken.

The manual began with 18 pages of introductory material that described the research objectives, the development of the TAT as a personnel assessment device, a description of the administration of the TAT, and the development of the cues and the manual, as well as instructions on how to use it. The final chapter of the introduction to the manual concerned the description of the 10 stimulus cards. Permission was obtained from the President and Fellows of Harvard College to reproduce reduced versions (about 4 x 5") of the cards that had been employed in the study. Permission was also secured from John Wiley & Son, Incorporated, and from Dr. William E. Henry to reproduce his descriptions of the cards as they appeared in the Analysis of Fantasy.

The main body of the manual consisted of 18 segments, one for each of the 18 variables to be rated. Each section began with a reproduction of the rating scale that the evaluator was to use. This was followed by two to two and one-half pages of discussion of the nature of the variable with special reference to its expression through the TAT. Each section then went on to describe the indicators from the protocol as a whole that related to each of the steps on the particular

scale under consideration. Following this was a section in which reference was made to specific cards in which clues to the variable being discussed were most likely to occur, and the nature of such clues.

This was followed by a discussion of the use of these clues in arriving at the rating. Next, examples from actual cases illustrating each step on the scale were presented and commented on. The final portion of each of the 18 parts of the manual consisted of a self-test in which from three to five protocols were reproduced and questions asked concerning their interpretations, with the answers provided for feedback purposes.

The final section of the manual presented a complete analysis of a TAT record. Information about the age, birth order, father's occupation, education, marital status, job sought, and time taken are given before the recording of the protocols to the 10 cards. Then, ratings for each of the 18 scales are given, together with the analyst's basis for making his rating on each of the 18 scales.

With the reproduction of the manual the first phase of the study supported by the research grant from the National Institute of Health, Research Grant M-2158, was completed. The second phase consisted of the determination of the reliabilities and factorial structure of the scales. This was jointly supported by the Personnel Research and Development Corporation and the Office of Naval Research.

### III. THE RESEARCH DESIGN

#### A. Statistical Design

The basic aim of the research with the system developed in the manual was to test the reliability of judgments based upon it. A second objective was to determine the factorial structure of the 18 scales.

The fundamental requirement for determining the reliabilities of the 18 scales is, of course, to have an adequate number of protocols, each rated by two evaluators. One possible approach would have been to have selected two raters and have each evaluate the same number of cases. To do this would have required each of the evaluators to have provided 18 ratings on a minimum of 100 cases. The evaluation of a case following the method prescribed in the manual would require between one and three hours after the evaluator had become completely familiar with the TEMP procedures. It would have been extremely difficult to have found even so few as two individuals with industrial personnel experience

who would have been willing to have undertaken this chore. Furthermore, as with all clinical or quasi-clinical techniques, the reliability (and, of course, the validity) is as much or more a function of the individual evaluator, his skill, perceptiveness and conscientiousness as it is of the evaluation technique employed. Using only two evaluators, chance alone would have been the chief determining factor in our results. Other contaminants, such as practice effect and task fatigue, would have undoubtedly influenced the results.

Fortunately, H. E. Brogden \* has demonstrated that the use of multiple evaluators in rating reliability and validity studies is mathematically equivalent to the use of just two raters for this purpose. On this basis, we decided to use 50 evaluators. Each evaluator would rate the protocol of five cases common to all 50 of them (the "common" cases). In addition, 125 cases would each be rated by two raters (the "uncommon" cases). Thus, each evaluator would first analyze the five common cases and, having had this experience in the use of the method, would then proceed with the remaining five cases, each of which would be evaluated by one other evaluator. This plan of data collection, using 50 raters who each evaluated 10 records, would provide 50 ratings on each scale for each of 50 cases, and two ratings on each scale for 125 cases. These data would enable us to compute two estimates of reliability for each scale -- one based on the five common cases, and the other on 125 uncommon cases. Both of these sets of data are amenable to the use of intra-class correlations, thus resolving the problem of the large number of raters in the first instance, and the problem of how to constitute the "x" and "y" variables in the second. Furthermore, we planned to check these reliability measures against Horst's more generalized formula for reliability determination (19).

It is relatively infrequent in the industrial situation that the reliability of rating scales is determined or published, whether they are used for criterion or personnel action situations. Where independence of ratings has been maintained, reported reliabilities have generally ranged from about .30 to about .60. Thorndike (30) mentions a study of Naval efficiency ratings where the correlations are in the .30's between successive ratings of officers who remained on the same post, but dropped to between zero and .20 when ratings made on shore duty are

(\*) Personal Communication

compared with those made on the same officers on sea duty. Our expectations, therefore, should not be too great for exceptional results in this study. It should also be remembered that a reliability of .30 could yield a validity of .55, which would be unusually high for predictive studies with projective techniques. We therefore made the arbitrary decision to retain for further research those scales whose reliability was .30 or higher.

#### B. The Experimental Data

The TAT protocols selected for this study were drawn from the assessment files of Personnel Research and Development Corporation. The 130 cases were selected by taking every odd-numbered case, beginning with No. 1, and proceeding up through about 320. In some instances, there was incomplete data or the TAT was not given, and these cases were rejected. The cases were selected before the development of the indicators for the manual, so that they would not be used in searching for such cues, a process which would have resulted in applying a scoring system back to the sample from which it was derived. To select the five common cases from the 130 selected, a table of random numbers was used.

The common cases were derived from four different companies, there being two cases from one company. The positions involved were: one sales candidate, one general administrative person at the middle level of management, two engineers, and one industrial engineering trained man who had an administrative job somewhat related to this work. This sample of five is quite representative of the total assessment population from which it was drawn, with the exception that there was one less sales candidate and one more engineer than would be proportionately true for the total population.

The 125 uncommon cases were drawn from 35 different companies. These companies include both large and small organizations representative of the electronics, railroad, insurance, metal working, clothing, banking, coal-mining, paint, and steel manufacturing industries. The three companies most heavily represented accounted for 52 of the 125 cases. The remaining 73 were sprinkled throughout the other 32 companies. Of the 125 common cases, 75 were men being considered for sales and sales management positions. The remaining 50 were management people, including engineering and related technical specialists, factory management, personnel and finance people. Fifty-one of the 125 cases were in a management capacity involving line responsibilities of some nature, either in sales or manufacturing.

The TAT protocols were typed exactly as in the original, hand-written protocols, so that the evaluators did not have access to the original material. Thus, they could not be influenced by such factors as handwriting and neatness. Each story was put on a separate page and the protocol was stapled together. An identification sheet was included with each protocol, a copy of which is included herein as Appendix B. This sheet provided the age, education, and marital status of the case, plus the time taken by him to complete the protocol (Slow, Fast, Average). An explanatory note was added so that evaluators would know that misspellings and omissions of words in the stories were exactly as had been written in the original protocol. No other information about the man, his job, etc. was provided to the evaluator. In effect, these were to be "blind analyses." Each evaluator was provided with an instruction sheet outlining his task and a set of rating scales upon which to indicate his ten sets of ratings. The instruction sheet is included as Appendix C.

#### C. The Evaluation Population

##### Description

In developing this research it was intended that the 50 evaluators be individuals with some experience and training in personnel work and/or personnel psychology, but that they not be experienced in the use of projective techniques. Since the contemplated technique was to be used by people with some familiarization with the industrial world and with the selection process, the aim was to obtain evaluators employed by corporations in personnel capacities, and for whom selection of managerial people was a major responsibility. In order to obtain the 50 evaluators, the project staff developed a list of acquaintances throughout the country who fitted this category, including both persons with a degree in psychology (either M.A. or Ph.D.) and those without this training, but all with experience and training in personnel selection. Requests for participation were forwarded to these people, and a basic group of 50 was selected from the larger number put together for this purpose. Since the design of the study made it imperative that 50 evaluators be used, and each of these for specified cases, the problem of obtaining the evaluator population proved to be the biggest stumbling block to the completion of this research. That is, people who agreed to participate in the evaluation of the ten protocols dropped out at various stages without completing the task, requiring solicitation of new evaluators.

In some instances, it was necessary to obtain a fourth or fifth volunteer for a given evaluation position, as the previous three or four volunteers dropped out. This, of course, resulted in substantial delay in obtaining the 50 evaluations.

The final evaluation population consisted of 46 men and four women. Furthermore, because of the difficulty in getting evaluators, it was necessary to obtain some people who had some experience with projective techniques, not necessarily the TAT, but in some instances including this device. When this became necessary we decided to compare the reliabilities for people with different levels of background. The final 50 evaluators were categorized into three groups as follows:

Group A - 13 people with no training in psychology or projective techniques. All were in personnel work for business corporations or similar large organizations.

Group B - 27 people with training in psychology or personnel work but not in projective techniques. Nine of these people had an M.A. degree, and 18 had a Ph.D. degree in psychology. Three of these were in consulting work, three were university professors, and the other 21 were in personnel positions with industrial organizations.

Group C - Ten people with degrees in psychology and some training and experience in the use of projective techniques. All but one person was working in an industrial organization or as a consultant in the area of personnel selection.

#### Time for Completion of Evaluations

Analysis of the evaluations completed by the 50 evaluators indicated a wide range of time taken for completion of the total task. The range went from one day, in the case of one person, to 13 months in the case of another. In the latter case, the last seven of the ten cases were completed in the last three weeks of the 13th month. Similarly, many people who completed the task in 60 days did the last six to eight cases in the last few days. In other words, where a relatively long period of time was taken to complete the task, there was a speeding up process in which the last several cases were completed in relatively short order. The following table presents the distribution of time taken to complete the task.

TABLE 2

Time taken to Complete Evaluations

<u>Length of Time taken</u>	<u>Number of Evaluators</u>
Seven months or more	4
Five to seven months	1
Three to five months	2
Two months	7
One month	6
Three weeks	8
Two weeks	6
One week	8
One to five days	4
No data	4
Total	<hr/> 50



It will be seen from Table 2 that 30 evaluators, or 60% of the group, completed their evaluations within a three-week period, and that all but 14 people completed it within a month. Since many of the 14 who completed the task in more than a sixty-day period indicated that they had done the last six or eight cases within a two or three-week period, it seems that once the evaluators got into the task, they did most of their work within a relatively short period of time. Most of the work was done within a two or three-week period following the studying of the manual and familiarization with the technique.

#### Relevant Comments by Evaluators

A number of the evaluators volunteered some comments about their experience in this study. From the letters received, the following observations have abstracted where they occurred more than once:

1. The first few cases took a lot of time, as much as five hours in one instance. Then, there was a speeding up process until the last few were completed more quickly. However, it was clear that, even with some practice and familiarization, it took almost two hours to complete the 18 ratings of one case.
2. Some evaluators felt that the last six or so scales were harder to evaluate than the first 12.
3. Differentiation between Scale 11, "Level of Ambition," and Scale 12, "Intensity of Ambition," proved to be difficult, and the value of attempting this differentiation was questioned.
4. Some protocols did not contain enough indicators for rating a given scale. This was raised in connection with the question of whether a judgment was warranted on that scale in instances such as this.
5. Some of the projectively-trained psychologists felt that, even though the time to evaluate a person was substantial with this technique, they still did not spend sufficient time on each of the scales to insure a truly high quality judgment.

The most detailed comments were furnished by one evaluator, a Ph.D. in psychology with some experience with projectives. Because of the centrality of his comments, the following excerpts from his letter are included:

- "1. In the first place, I found this scoring system to be very complex and detailed. Besides the large amount of time and preparation which it requires, there is another objection: for the different qualities it is almost impossible to keep in mind the five indicator-categories and to read the stories. Nor is it possible, on the other side, to keep in mind the stories and to read the indicators. You have to break down the totality and to do one of the following things:
  - a. To evaluate story by story. But the difficulty here is that there are a lot of indicators (not only formal indicators but also aspects of the content of the stories themselves) with which you have to judge the protocol as a whole.
  - b. To judge scorings category by category. But then you lose the overview and the comparisons which are necessary for careful consideration of the categories.
- "2. In the second place, the indicators refer to formal analysis as well as content analysis, plus a personal clinical, intuitive impression. This sometimes required me to make a quick switch in concentration and attitude. In addition, it can happen that the careful, detail analysis does not agree with the general feeling. It depends on the personal ideas of the evaluator which of the two he stresses most.
- "3. Some of the 18 qualities have a rather pure continuum from A to E. It is just a unidimensional 5-point scale for judgment of a personality trait or a certain performance. But I felt other categories do not have this clear continuum. Here the five levels are not quantitatively different, but qualitatively. This implies, to my opinion, two difficulties with the scoring:"

- "a. For example there exist some indications which point in the direction of an E score, but these indications are not quite strong enough to justify this E scoring. You would like to score one level lower (D), but then it turns out that this category is qualitatively-different and does not fit at all. These situations happen especially with the qualities: 7 (involvement in work), 10 (reaction to failure and frustration), 13 (relation with superiors), 14 (organizing and directing others), and 16 (insight in others).
- b. In connection with the above I sometimes found it necessary to choose not the best fitting category but the 'less worst' category. Maybe it is worthwhile to consider using a category such as 'none of these,' because of the possibility that the protocol withdraws itself almost completely from the quality in question. Another possibility would be the use of more specific examples in the 'Using the Indicators' section.
- c. When there is not a continuum, but when the choice possibilities are qualitatively different, these can only be taken on the basis of an underlying theory. And as a theory it is open to criticism. For instance, in my opinion, the five possibilities mentioned as a reaction to frustration (10) do not cover the whole range of possible reactions."

#### IV. RESULTS

##### A. Reliability Coefficients

The intra-class correlations and "Horst" reliabilities are given separately for the common and uncommon cases in Table 3. In 11 out of the 18 scales, the uncommon cases had a higher intra-class correlation than did the common cases. In two cases they were identical, and in the other five the common cases had a higher reliability than did the uncommon cases. Reliabilities computed by the

TABLE 3  
Basic Reliability Indices for Common and Uncommon Cases (decimals omitted)

		Intraclass r	Horst's Reliability	Spearman Brown 50 raters	Intraclass r	Horst's Reliability	Spearman Brown 2 raters
1.0	Accuracy of Perceptual Interpretation	17	88	91	28	43	44
2.0	Amount of Productivity	44	96	98	55	69	71
3.0	Consistency of Productivity	02	42	50	18	30	31
4.0	Whole vs. Detail Approach to Problems	50	98	98	10	18	18
5.0	Originality of Expression	61	98	99	61	75	76
6.0	Emotional Control	31	95	96	42	59	59
7.0	Involvement in Work	58	98	99	44	59	61
8.0	Need for Clear Assignment and Direction	08	77	81	21	34	34
9.0	Need for Emotional Support in Work	31	94	96	33	49	49
10.0	Reaction to Frustration and Failure	15	87	90	15	26	26
11.0	Level of Ambition	32	94	96	48	64	65
12.0	Intensity of Ambition	22	90	93	24	37	39
13.0	Relations with Superiors	27	94	95	04	07	07
14.0	Ability to Organize and Direct Others	19	89	92	34	52	51
15.0	Consideration for Subordinates	39	96	97	25	40	40
16.0	Insight into Others	26	93	95	40	55	58
17.0	Social Skills	48	97	98	36	53	53
18.0	Social Dominance	27	93	95	33	49	50

Horst method were universally higher than those by the intra-class method and quite substantially so for the 50 common cases. Review of Horst's article (19) led us to the belief that, as applied to these data, Horst's formula was equivalent to having applied a Spearman-Brown correction to the intra-class correlations. To verify this hypothesis, the Spearman-Brown formula was applied to each of the two sets of correlations, raising the number of raters for the common cases to 50, and for the uncommon cases to two. As may be seen from Table 3, this turned out to be almost exactly true, except where the intra-class correlations are extremely low. Since in the rating situation the Spearman-Brown yields the reliability for an average of "n" raters and it is highly unlikely in analyzing written protocols to the TAT that two or more independent evaluations will be made and averaged, it is the basic intra-class correlations rather than the Horst or those corrected by the Spearman-Brown prophecy formula that must apply. These, as may be seen from Table 3, range in the instance of the common cases from .02 to .61, and in the uncommon cases from .03 to .61. For the uncommon cases, ten of the 18 scales met our criterion of having a reliability in excess of .30. Nine of the 18 reliabilities on the five common cases exceeded .30. Only these ten scales, plus "Consideration for Subordinates" (whose reliability in revised form was .36), were retained for the revised manual.

While the product moment, or in this case intra-class correlation, between raters provides an abstract measure of reliability, the same data can also be viewed in terms of the degree of disagreement among raters. These are presented for the 125 uncommon cases in terms of percentages in Table 4. From here, it may be seen that, with a few exceptions, disagreement of more than one step occurred in less than 20% of the pairs of ratings. Thus, from a practical, as well as a statistical point of view, it may be seen that the selected scales may be used with confidence.

In the five common cases, the 50 raters would yield 1,225 paired comparisons for each of the 18 scales. Rather than presenting the degree of disagreement for these cases, therefore, Table 5 presents the frequency distribution of the 50 ratings of each individual on each retained scale. Examining this table, we see there is a wide variation of the distribution, not only from case to case, but from scale to scale as well. On Scale 5, "Originality of Expression," for example, the modal rating for Case 126 is in category 2 and contains only 19 cases, or 38% of the ratings. The adjacent categories 1 and 3 contain 17 and 3,

TABLE 4  
Percent Disagreement by Scale, 125 Uncommon Cases  
Data Calculated from Scatter Plots

Units of Disagreement	0	1	2	3	4
1.0 Accuracy of Perceptual Interpretation	59.2	20.8	16.8	3.2	0.0
2.0 Amount of Productivity	44.8	48.8	6.4	0.0	0.0
3.0 Consistency of Productivity	64.8	20.8	9.6	4.8	
4.0 Whole vs. Detail Approach to Problems	42.4	38.4	15.2	4.0	0.0
5.0 Originality of Expression	46.4	41.6	9.6	.8	1.6
6.0 Emotional Control	45.6	39.2	12.8	1.6	.8
7.0 Involvement in Work	38.4	45.6	14.4	.8	.8
8.0 Need for Clear Assignment and Direction	32.0	42.4	16.0	6.4	3.2
9.0 Need for Emotional Support in Work	38.4	41.6	15.2	4.0	.8
10.0 Reaction to Frustration and Failure	42.4	29.6	17.6	6.4	4.0
11.0 Level of Ambition	44.0	37.6	15.2	2.4	.8
12.0 Intensity of Ambition	35.2	40.0	19.2	4.0	1.6
13.0 Relations with Superiors	35.2	36.8	21.6	5.6	.8
14.0 Ability to Organize and Direct Others	50.4	35.2	12.8	1.6	0.0
15.0 Consideration for Subordinates	42.4	39.2	12.0	5.6	.8
16.0 Insight into Others	43.2	37.6	16.8	2.4	0.0
17.0 Social Skills	47.2	40.8	8.8	.8	2.4
18.0 Social Dominance	43.2	46.4	8.0	2.4	

- 0 - Both ratings same
- 1 - Two ratings in adjacent categories
- 2 - One intervening category between ratings
- 3 - Two intervening categories between ratings
- 4 - Three intervening categories between ratings

TABLE 5

Distribution of Ratings by Retained Scale  
for the Five Cases in the Common Population

Scale	Case	R A T I N G   C A T E G O R Y					
		1	2	3	4	5	6
2.0 Amount of Productivity	126	0	3	22	23	2	
	127	1	1	5	40	3	
	128	13	25	11	0	1	
	129	12	1	3	13	21	
	130	4	21	25	0	0	
5.0 Originality of Expression	126	3	19	17	4	5	2
	127	0	3	6	34	6	1
	128	41	7	0	0	1	1
	129	5	4	0	7	24	10
	130	27	14	7	0	1	1
6.0 Emotional Control	126	3	12	22	5	8	
	127	0	0	9	28	13	
	128	25	18	0	4	3	
	129	16	10	6	7	11	
	130	8	18	11	7	6	
7.0 Involvement in Work	126	8	23	16	1	2	
	127	0	3	13	33	1	
	128	14	33	2	0	1	
	129	1	17	4	14	11	
	130	29	15	6	0	0	
9.0 Need for Emotional Support in Work	126	4	12	20	13	1	
	127	0	10	15	24	1	
	128	30	14	2	2	2	
	129	21	10	5	9	6	
	130	2	6	24	7	11	
11.0 Level of Ambition	126	29	11	7	3	0	
	127	2	16	20	12	0	
	128	25	19	4	2	0	
	129	6	23	8	13	0	
	130	22	15	12	1	0	
14.0 Ability to Organize and Direct Others	126	33	17				
	127	7	43				
	128	42	8				
	129	26	24				
	130	41	9				

TABLE 5  
(Continued)

Scale	Case	R A T I N G   C A T E G O R Y					
		1	2	3	4	5	6
15.0 Consideration	126	4	41	5			
	127	1	12	37			
	128	3	45	2			
	129	3	33	14			
	130	9	37	4			
16.0 Insight into Others	126	1	8	29	6	6	
	127	0	4	9	30	7	
	128	18	17	10	1	4	
	129	3	18	0	10	19	
	130	8	12	22	2	6	
17.0 Social Skills	126	0	13	22	8	0	7
	127	0	2	7	27	10	4
	128	8	35	7	0	0	0
	129	0	12	9	7	0	22
	130	7	34	7	1	0	1
18.0 Social Dominance	126	7	22	14	7		
	127	1	26	19	4		
	128	36	9	2	3		
	129	28	12	1	9		
	130	9	13	10	18		



respectively, for 34 and 6% of the ratings, giving a total in the modal and two adjacent categories of 78% of the ratings. Twenty-two percent of the ratings differed by more than one category from the mode. Forty-one, or 82% of raters, however, rate Case 128 as being in category 1 on this scale, and only 4% deviate by more than one rating category from the mode. This, incidentally, is the scale with the highest reliability in both the common and uncommon cases computations.

While we had hypothesized that the reliabilities of the common cases would be lower than those of the uncommon because of the practice effect, a second factor that may have contributed to this phenomenon is the fact that there were only five protocols in the common cases, as opposed to 125 in the uncommon cases. Thus, it might be expected that the uncommon cases would present a larger between-case variability than the common cases. To the extent that such restriction in range existed among the five common cases, it would be expected that their reliabilities would be lower.

#### B. Analysis of Scatter Plots

The steps on each of the 18 scales had, of course, been constructed on an a priori basis by members of the project staff, with extensive experience in the evaluation of protocols from similar populations. The intent in the preparation of the step on each scale was to write a linear scale with behaviorally distinguishable steps. That the scale writers would not be wholly successful in accomplishing these objectives is a foreborne conclusion and is borne out by the comments of some of the evaluators. In order to make such corrections as were possible on an empirical basis, the 18 scatter plots for the 125 uncommon cases were prepared. These were examined in the light of marginal distributions as well as of the scale content. Using these as guidelines, two of the 10 scales whose reliabilities were in excess of .30 were revised. These were Scale 7, "Involvement in Work," and Scale 14, "Ability to Organize and Direct Others." A third scale, "Consideration for Subordinates," which had an original reliability of .25, yielded a reliability of .36 when reduced from a five to a three-point scale. This became the 11th scale to be retained. Details of these changes are described in the section entitled "Revision of Manual" below.

The numbers assigned to the scale steps are, of course, arbitrary. While it was the intent of the project personnel who designed the scales to construct

linear continua in which the differences between adjacent steps would be quantitative rather than qualitative, any attempt to have pre-judged the distribution or to have hoped that the resulting distributions would have been recti-linear, or that they would have been normal, would have been naive. For the uncommon cases the relative frequency distributions of the 11 scales with reliabilities above .30 are given in Table 6. It may be seen that while there is a general tendency for these 11 distributions to follow some sort of bell-shaped distribution, none resembles a normal curve. For those wishing to use the scales quantitatively rather than descriptively and are willing to assume that the characteristics underlying each of the scales is normally distributed in the population tested, the standard score equivalent of the mid-point of each of the scale steps for each of the 11 surviving variables is also presented in Table 6.

#### C. Reliability by Rater Type

As has already been mentioned, our attempt to secure a homogeneous population of raters had to be abandoned because of the difficulty in securing evaluators. Our final evaluation population could be divided into three groups. Group A consisted of 13 raters who were employed in personnel work but had no training in either psychology (beyond what they may have had in undergraduate courses) or in the analysis of projectives. Group B consisted of 27 personnel men trained in psychology to either the master's or doctor's level, but with no special training in the use of projective devices. Group C consisted of ten psychologists beyond the master's level who were trained in the use of projective techniques, including the TAT. Table 7 presents the intra-class correlations by these rater breakdowns for the five common cases. For comparison purposes, the intra-class correlations resulting from the use of all 50 cases is also presented. There is no consistency in the table. In all 18 scales, one or more of the sub-populations of evaluators was more reliable than the total group. In seven scales, the untrained raters yielded the highest reliability. In five scales, psychologists without projective training yielded the highest reliability; while in six scales the trained projective analysts were highest. With several noticeable exceptions, however, the general magnitude of the correlations is fairly consistent. This is particularly true of those scales with practical reliability. For example, Scale 3 shows an intra-class correlation for the total

TABLE 6  
Relative Frequency and Standard Score Values  
of the Retained Scales (as revised)

		A	B	C	D	E	F
2.0	Amount of Productivity	% 6.0	25.2	43.2	24.0	1.6	
	SS	-1.88	-.89	.07	1.14	2.41	
5.0	Originality of Expression	% 26.4	31.2	20.0	17.2	5.2	
	SS	-1.12	-.20	.46	1.09	1.94	
6.0	Emotional Control	% 12.8	28.4	40.8	13.6	4.4	
	SS	-1.52	-.61	.30	1.22	2.01	
7.0	Involvement in Work	% 18.0	30.6	30.2	21.0		
	SS	-1.34	-.43	.36	1.25		
9.0	Need for Emotional Support in Work	% 8.0	23.2	34.4	26.8	7.6	
	SS	-1.75	-.86	.04	.81	1.77	
11.0	Level of Ambition	% 21.2	34.8	23.2	16.4	4.4	
	SS	-1.25	-.29	.46	1.45	2.01	
14.0	Ability to Organize and Direct Others	% 18.9	30.6	50.5			
	SS	-1.31	-.41	.67			
15.0	Consideration for Subordinates	% 4.8	51.6	43.6			
	SS	-1.98	-.51	.78			
16.0	Insight into Others	% 14.4	10.8	44.0	24.4	2.4	
	SS	-1.46	-.85	.03	1.05	2.26	
17.0	Social Skills	% .8	25.6	42.8	22.4	6.0	2.4
	SS	-2.65	-1.12	.06	.86	1.61	2.26
18.0	Social Dominance	% 27.6	46.8	18.4	7.2		
	SS	-1.09	.03	.98	1.80		

TABLE 7

Intraclass Correlation by Rater Type  
Common Cases - 5 Cases Rated by all Raters  
(decimals omitted)

		A 13 raters <u>R<sub>13</sub></u>	B 27 raters <u>R<sub>27</sub></u>	C 10 raters <u>R<sub>10</sub></u>	D All raters
	<u>R<sub>50</sub></u>				
1.0 Accuracy of Perceptual Interpretation	17	12	<u>24</u>	04	28
2.0 Amount of Productivity	44	<u>62</u>	38	39	55
3.0 Consistency of Productivity	02	<u>05</u>	01	-04	18
4.0 Whole vs. Detail Approach to Problems	50	<u>59</u>	45	56	10
5.0 Originality of Expression	61	<u>68</u>	60	49	61
6.0 Emotional Control	31	37	28	<u>38</u>	42
7.0 Involvement in Work	58	52	<u>64</u>	48	44
8.0 Need for Clear Assignment and Direction	08	<u>08</u>	08	04	21
9.0 Need for Emotional Support in Work	31	11	<u>36</u>	31	33
10.0 Reaction to Frustration and Failure	15	02	16	<u>30</u>	15
11.0 Level of Ambition	32	14	37	<u>49</u>	48
12.0 Intensity of Ambition	22	<u>35</u>	19	12	24
13.0 Relations with Superiors	27	24	<u>29</u>	23	04
14.0 Ability to Organize and Direct Others	19	08	16	<u>26</u>	34
15.0 Consideration for Subordinates	39	26	<u>47</u>	25	25
16.0 Insight into Others	26	12	33	<u>37</u>	40
17.0 Social Skills	48	49	44	<u>51</u>	37
18.0 Social Dominance	27	<u>34</u>	20	33	33

A - No training in psychology or projectives  
B - Training in psychology but not projectives  
C - Experienced in projectives

group of .02, and for groups A, B, and C, .05, .01, and -.04, respectively. Similarly, for Scale 5, "Originality of Expression," the total group yielded an intra-class correlation of .61, and the three sub-groups .68, .60, and .49, respectively.

The same procedure was attempted for the uncommon cases. Among the 13 untrained raters, only five rated the same cases as other untrained raters. Twenty-five of the 27 non-projective psychologists rated identical cases, but only five of the ten trained projective analysts had cases in common. This accounts for 35 of the 125 uncommon cases, leaving 90 in which the raters belonged to different groups. These data are presented in Table 8. Here, there is even less consistency than was present among common cases. Again, in each of the 18 instances, one or another of the homogeneous groups yielded a higher reliability than the total group; and in all 18 scales, one or more of the homogeneous groups demonstrated a higher reliability than was demonstrated by the 90 cases where the two raters belonged to different groups. Variability is much more extensive here than was true for the common cases, but it should also be noted that the magnitude of correlations of the individual groups is frequently considerably higher than that of the total group. In six of the 18 scales, the untrained raters demonstrated the highest reliability. In five, the psychologists without projective training showed the highest reliability, while the projective analysts showed the highest reliability in seven of the 18 scales. In reviewing this table, the reader should bear in mind that both the untrained group and the projective analysts consisted of only five raters. The greater variability may be purely a function of the number of cases in the respective rater populations.

#### D. Intercorrelations and Factor Analysis

Table 9 presents the matrix for the 500 ratings rendered by the 50 raters on the 18 original scales. The highest correlation in the matrix is .70 between "Intensity of Ambition" and "Level of Ambition." The next highest correlation in the matrix is .51 between "Social Skills" and "Originality of Expression," and between "Originality of Expression" and "Involvement in Work." While there are some correlations in the .30's and .40's, many of the relationships are considerably below this level. A centroid factor analysis yielded only two factors, the second of which was reduced to zero by a simple orthogonal rotation.

TABLE 8

Intraclass Correlation by Rater Type  
Uncommon Cases - Each Case Rated by Two Raters

		A 13 raters n=5	B 27 raters n=25	C 10 raters n=5	Het.* n=90
1.0 Accuracy of Perceptual Interpretation	28	40	<u>50</u>	00	23
2.0 Amount of Productivity	55	68	43	<u>73</u>	54
3.0 Consistency of Productivity	18	<u>80</u>	25	60	13
4.0 Whole vs. Detail Approach to Problems	10	<u>70</u>	56	-100	-06
5.0 Originality of Expression	61	<u>71</u>	59	<u>71</u>	60
6.0 Emotional Control	42	21	07	<u>83</u>	47
7.0 Involvement in Work	44	60	30	<u>66</u>	46
8.0 Need for Clear Assignment and Direction	21	54	31	<u>87</u>	15
9.0 Need for Emotional Support in Work	33	12	<u>60</u>	48	27
10.0 Reaction to Frustration and Failure	15	<u>73</u>	12	-60	16
11.0 Level of Ambition	48	48	<u>51</u>	29	48
12.0 Intensity of Ambition	24	22	38	<u>68</u>	21
13.0 Relations with Superiors	04	-20	<u>44</u>	43	-04
14.0 Ability to Organize and Direct Others	34	30	<u>40</u>	32	33
15.0 Consideration for Subordinates	25	<u>90</u>	42	28	16
16.0 Insight into Others	40	45	38	<u>83</u>	44
17.0 Social Skills	36	<u>87</u>	60	10	30
18.0 Social Dominance	33	40	45	<u>83</u>	29

A - No training in psychology or projectives

B - Training in psychology but not projectives

C - Experienced in projectives

\* Heterogeneous cases. Cases which were rated by two raters who were not of the same type.

1.0 Accuracy of Per19	01	-24	-06	-01	06	13	15	00	00	09	-09	14	04	04	09	01
2.0 Amt. of Productivity-05	08	49	01	26	50	20	23	26	30	33	-02	32	22	35	41	16
3.0 Consistency of Prod.	-03	-03	30	11	02	18	-06	03	-03	-01	-04	01	-05	-02	-06	02
4.0 Whole vs. Detail Approach					14	04	08	17	16	15	10	09	18	13	20	09
5.0 Originality of Expression				48	51	33	16	24	31	27	02	15	31	39	51	20
6.0 Emotional Control					32	33	29	15	20	20	-13	09	17	33	34	30
7.0 Involvement in Work						28	24	29	43	40	-01	38	33	38	48	17
8.0 Need for Clear Assign. and Direction							31	20	19	20	-18	10	08	2	26	28
9.0 Need for Emotional Support in Work								27	24	24	-29	22	12	16	26	38
10.0 Reaction to Frustration and Failure									16	16	-05	23	22	26	27	15
11.0 Level of Ambition										70	-05	37	20	22	31	23
12.0 Intensity of Ambition											-05	35	21	24	28	28
13.0 Relations with Superiors												-03	17	-11	-02	-31
14.0 Ability to Organize and Direct Others													14	21	31	24
15.0 Consideration for Subordinates														38	31	-07
16.0 Insight into Others															38	17
17.0 Social Skills																
18.0 Social Dominance																

TABLE 9

Intraclass Correlation  
Between Rating Scales

N = 250 (50 raters - 5 uncommon cases  
5 common cases)

Table 10 presents the results of the factor analysis. All of the retained scales have substantial loadings on this factor, ranging from .39 for "Consideration for Subordinates" to .68 (the highest loading on the factor) for "Involvement in Work." Looking at the five retained scales which have a loading of .60 or better, we find "Amount of Productivity," "Originality of Expression," "Involvement in Work," "Level of Ambition," and "Social Skills."

The magnitude of the correlations in the matrix as a whole is encouraging. Compared with the level of correlations obtained from rating scales based on either selection interviews or performance appraisal, there is a noticeable absence of the all too common halo. In no case is as much as half of the variance of a given scale explained by the other scales in the matrix.

#### V. REVISION OF MANUAL

The 11 scales which remain in the revised manual were reviewed for revisions within each. For all 11, a routine editing was performed to enhance the clarity and simplicity of the presentation. This included the minor change of combining indicators from the total protocol with indicators from the individual stories, as it was found that this was an unnecessary and cumbersome breakdown. For seven of the scales, this was the only revision performed.

For Scale 9, "Need for Emotional Support in Work," some of the indicators from Scale 10, "Reaction to Failure and Frustration," were included where they seemed to be appropriate. This was indicated because of the relatively high correlation between the ratings for these scales and the fact that at least two of the categories for each of these scales were highly repetitious. Since "Reaction to Failure and Frustration" was dropped in the revision, it was felt that this change would make "Need for Emotional Support in Work" a more substantial scale by including some indicators that would clarify the scale steps.

Three of the scales required some revision as to scale categories. For Scale 7, "Involvement in Work," the "E" category was dropped. There were no ratings placed in this category, and it was felt that four rating steps posed about as meaningful a discrimination as evaluators could comfortably deal with. Similarly, Scale 14, "Ability to Organize and Direct Others," and Scale 15, "Consideration for Subordinates," were reduced to three categories. For Scale 14, categories "D" and "E" were eliminated. It was felt that "C" and "E"



TABLE 10

Original and Rotated Orthogonal  
Factors of 18 Rating Scales in TEMP

	$K_1$	$K_2$	$K_1'$	$K_2'$	$n_K^2$
1.0 Accuracy of Perceptual Interpretation	12	05	12	04	02
2.0 Amount of Productivity	60	04	60	-02	36
3.0 Consistency of Productivity	-02	00	-02	00	00
4.0 Whole vs. Detail Approach to Problems	24	02	24	-01	06
5.0 Originality of Expression	63	09	64	02	41
6.0 Emotional Control	52	04	52	-02	27
7.0 Involvement in Work	67	11	68	04	46
8.0 Need for Clear Assignment and Direction	44	07	44	03	20
9.0 Need for Emotional Support in Work	50	12	51	06	27
10.0 Reaction to Frustration and Failure	41	03	41	-02	17
11.0 Level of Ambition	59	04	60	-02	35
12.0 Intensity of Ambition	60	05	60	-01	36
13.0 Relations with Superiors	-16	01	-16	-01	03
14.0 Ability to Organize and Direct Others	46	03	46	-02	22
15.0 Consideration for Subordinates	38	03	39	-01	15
16.0 Insight into Others	54	12	55	06	31
17.0 Social Skills	65	07	66	00	43
18.0 Social Dominance	44	08	45	04	20

were closely related, as were "B" and "D". Since there were few ratings in categories "D" and "E" the scale was simplified in this way. This change increased the reliability of the scale from .34 to .41.

For Scale 15, analysis of the scatter plots had indicated a great deal of confusion in the minds of the raters between categories "B," "C," and "E." For this reason, a new scale was developed which retained category "D" intact, but which made changes in categories "A" and "C" so that the new three-point scale would be more meaningful behaviorally as well as practically. The revision of this scale increased its reliability from .25 to .36 and made it eligible for inclusion in the operating edition of the manual.

In the revised manual, the scales are presented in a new order. This was done so as to introduce some of the easiest scales first, as well as to enable evaluators to deal initially with rating scales which would be helpful with later scales. The new order starts with considerations of productivity and quality in work, moves to factors of ambition and emotional stability, and then into social relationships. The final two scales, having to do with supervision of people, are more complex and inferential in their nature, and it was felt that putting these last would make them easier to evaluate when once having made an evaluation of the other nine areas. The new order and reliability of the scales in the revised manual are presented in Table 11, together with the original number of the scale.

In addition to revision of the scales as such, several other changes were made in the manual. The complete analysis of a record was revised to conform with the shortening of the scales, and also to bring about some greater clarity in writing. Finally, a section was prepared on uses of the revised manual. This section follows next in this report.

## VI. USES OF THE MANUAL

### A. For Training

This manual was designed primarily to train people in personnel work who have some responsibility for selection, but who lack professional training in projective techniques. It is our hope that the manual will enable them to extract some significant conclusions from the written record, even though they may not make a comprehensive analysis. The manual is intended to stand alone, and the

TABLE 11

Order and Reliability of Scales in Revised Manual

	<u>Original Number</u>	<u>Reliability</u>
1. Amount of Productivity	2	.55
2. Originality of Expression	5	.61
3. Involvement in Work (Revised)	7	.46
4. Need for Emotional Support in Work	9	.33
5. Level of Ambition	11	.48
6. Emotional Control	6	.42
7. Insight into Others	16	.40
8. Social Skills	17	.36
9. Social Dominance	18	.33
10. Ability to Organize and Direct Others (Revised)	14	.41
11. Consideration for Subordinates (Revised)	15	.36

examples have been chosen to cover a wide range of possibilities. The use of self-tests provides the individual learner with some opportunity to check on his understanding and skill.

Probably the manual is equally valuable as an instructional tool for professional psychologists who have not worked with the written version of TAT stories or who have not been exposed to stories told by a sample of managerial and sales people. The range of stories presented and discussed here also provides psychologists who work in clinics or in public schools with a new base line. Psychologists who have had no experience with projective techniques will find that the manual offers an easy introduction to the techniques of interpretation, even though they wish to develop a more comprehensive and dynamically-oriented interpretation eventually. The manual should also prove to be useful in training graduate students in psychology. For one thing, it contains a fair amount of normative information about the characteristics of the stimulus cards and the stories that are commonly told. The scale "Originality of Expression" would be particularly useful as a means of introducing the TAT, since the rationale behind the indicators is obvious and the indicators are not difficult to use. A more advanced exercise would require students to study the indicators for a scale and try to develop a rationale for each indicator. Scales such as "Emotional Control" or "Insight into Others" would be quite useful for this purpose.

Although some teachers may feel that the scales reflect too much of the "fixed trait" thinking about personality, they do have the advantage of requiring students to commit themselves to some definite conclusions. Many students of projective techniques use "weasel words" or hedge their statements so much that almost all meaning is drained out of their reports. The use of rating scales has at least the merit of avoiding this kind of escape. In addition, the stories presented in the manual provide considerable opportunity to perceive distortions, omissions, clarity of organization, and variations on common themes. All of these are valuable no matter what kind of interpretive scheme is used.

Since much published material on the TAT deals with pathology, it seems to us worthwhile to have a fund of stories available that come from a "normal" sample. It is well-known that students in training have a tendency to read more pathology into stories than is warranted. Exposure to the records of our sample may well diminish this tendency.

## B. In Personnel Assessment

The written version of the TAT has a number of advantages in personnel assessment. Like the oral TAT, it has the merit of permitting inferences about disguised or concealed aspects of motives and attitudes. It places the candidate in a position where he does not know how to present the most favorable impression, and consequently it bypasses the expected facade in the interview. The written version has some additional advantages, however. For one thing, it provides a permanent and accurate record of what the candidate actually did. And it does not require the time of a professional examiner to obtain this record. It can also be considered a "work sample" of the candidate's performance in an unstructured and potentially fatiguing situation.

The scales presented in this manual have been chosen with an eye to the qualities and skills needed by sales and management people. It is not necessary to use all of them in any particular case. In fact, it may often turn out that only two or three scales will be used, primarily to check on points that were not clearly revealed by other methods.

Combining data from an interview with the written stories is an absolute necessity in using the TAT. In our experience, the candidate ordinarily makes a somewhat better impression in an interview than he does on the TAT, and this difference must be considered in making the final evaluation. That does not always mean that the interviewer's judgment is the correct one.

Often, the inferences derived from the TAT cannot be checked until the man has been on the job for some months. We have taken the position that when unfavorable trends are discovered in the TAT stories, they may not reveal themselves directly in behavior until the man has settled comfortably into the job and has relaxed his guard a little. For example, suppose that the TAT interpreter believes that an applicant has the tendency to be hasty and slipshod in his work. When the applicant is hired, he wants to make an initially favorable impression and may take more than usual pains with his work. But as time goes on and he feels more secure in the new job, he may show the slipshod quality revealed in the TAT.

Interpreters using the TAT can do a better job if they are furnished with some information in addition to the written record. The very least that an interpreter should know is the age, sex, educational level, and marital status of the applicant. And, of course, he ought to have a good description of the

psychological qualifications desired for the position that is being filled. Additional data that are useful are the following: father's occupation, or some other indicator of the social class origins of the applicant; whether either of the parents died during the applicant's childhood or early adolescence, or whether a divorce or separation occurred; the sibling position of the applicant; some indication of the average grades in high school or college; a brief resume of the jobs held by the applicant since finishing his formal schooling. This kind of data provides a framework which enables more accurate inferences to be drawn from the record.

It is unwise to use the TAT for inferences about characteristics which can be measured more accurately and directly by other methods. For example, although it is possible to estimate the intelligence level of an applicant from his vocabulary in the TAT stories, a standard intelligence test is usually preferable. Inferences can also be drawn from the TAT about the kind of impact and manner that the applicant will show in face-to-face situations, but these characteristics are better observed in the interview. On the other hand, the written TAT stories present a somewhat better work sample than does the interview. And it can yield a more accurate picture of the individual's values. The written TAT should be regarded as a supplement to methods ordinarily used, not as a substitute for them.

### C. In Counseling

In vocational or career counseling, the TAT can provide the counselor with the same kind of data that it gives the selection specialist. He is thus enabled to make a somewhat better diagnosis of the problem. But the TAT can be used directly in the counseling process itself if the counselor is professionally trained. It is often a surprising and insight-producing experience for counselees to hear their stories read back to them some time after they have been written. The counselor may want to review the common stories given to each card before he reads the counselee's own stories.

The written stories can serve as a point of departure for much useful discussion. For example, one story may be definitely shorter than the others and filled with more erasures and misspellings. Inquiry into the counselee's feelings about the stimulus picture and his feelings while he was writing the story may generate some insights into his relations with others or his anxieties.

It is useful also to contrast the stories told early in the record with those told later, or to contrast stories to cards where both sexes are pictured with stories told to cards where only one sex is pictured. Of course, with some people, discussion of the reasons for writing certain kinds of stories may provoke defensiveness, justification, and superficial reasons just as in any other technique used in counseling.

#### D. In Research

Industrial psychologists and personnel specialists often wish to compare characteristics of two contrasting groups: for example, successful versus unsuccessful managers, or long tenure versus short tenure salesmen. Since written TAT records are rather easily obtained and permanently filed, a fund of usable information can be collected without much difficulty. Some or all of the scales presented in this manual can be used to score the records, and the data thus accumulated can be analyzed to show differences between the contrasting groups. Validity studies relating specific scales to success on the job are also quite possible.

One interesting area to explore with the written TAT is the predictive power of the scale "Originality of Expression." While it seems fairly clear that this scale does have predictive value in estimating the imaginativeness of people with a high degree of verbal skill, it is still quite unclear as to whether originality in the stories is connected with imagination in technical and scientific work. Comparisons between productive and unproductive research workers could teach us much about the TAT.

### VII. FURTHER RESEARCH

Although TEMP has uses in training and counseling, as well as in personnel research, our objective in designing the technique was to provide another predictor instrument which would have wide applicability in the selection of sales, technical, and managerial personnel. No predictor, not even one with demonstrated construct validity, can be assumed to be valid in any specific situation. The 11 scales in TEMP are no exception. At this stage in their development we know only that they are measures of moderate reliability, and that individuals with various degrees of technical training can learn to use the technique to make evaluations of written protocols to selected cards of the TAT. Whether any or all of such

ratings are valid predictors of job success remains at this point purely hypothetical.

In the realm of personnel selection, validation is a theoretical must. Yet, by and large, the basis on which most selection decisions are made remain unvalidated and often unvalidatable. Where objective tests of aptitude, proficiency, or personality are used, APA ethics are (unsuccessfully) insistent on empirical validation.

Selectors using a clinical approach, whether it be a face-to-face interview or a projective technique or a combination of these and other clinical devices, are assumed on the basis of their training and experience to be valid predictors of job success. It is rare that such predictions are empirically tested. In the literature of validation we infrequently find studies designed to measure the validity of a selection interview and, even more rarely, reports of the validity of projective techniques.

In the validation of an objective predictor, the resulting correlation expresses the relationship between the instrument and the criteria. Where clinical techniques, including this one, are concerned, validity is a function of both the device employed and its interpretation by whomever may be applying the technique. Thus, while the validity of an objective instrument may be specific for a given job, the validity of a clinical prediction is specific not only for a given job but for a given evaluator as well. Thus, with TEMP as with any other clinical technique, it is not possible to say with any degree of certainty, even after the usual type of validation study, that the technique is valid for a given job. Positive results would not enable us to generalize beyond the use of the technique by the particular interpreter who made the evaluations.

Projective techniques have generally been considered to be diagnostic tools primarily designed for the identification of pathologies. Consequently, most research employing these devices have been conducted in a clinical rather than industrial setting. It is our hope that in creating TEMP and making it available for use, we will stimulate the conduct of research on the use of the TAT in the industrial setting with normal populations. We hope that until more is known about its validity that TEMP will be used experimentally, and with all of the necessary precautions to determine the variability of its validity by different interpreters and the applicability of its scales to different kinds of sales,



technical, and managerial positions. Further data with respect to its reliability needs to be gathered within more homogeneous situations. It is also desirable to determine whether more intensive training in the technique would improve reliability. In the present study, the evaluators taught themselves, with only the manual to guide them.

We know that a certain proportion of over-zealous individuals will use the technique prematurely for operational purposes. We could, of course, withhold the technique from general publication until extensive additional research on it had been performed. But our facilities for conducting such research are limited. We make it available at this time in the hope that others will see in it the research possibilities that we do and add their facilities to ours, performing additional research studies which might be otherwise delayed for years or not at all conducted. We weigh the possible misuse of TEMP as the price that will need to be paid for making it available to others as a research device. It may be expected that the premature use of TEMP on a "faith validity" basis will result in a certain number of selection errors. But it is hardly likely that any more selection damage will be done by the inappropriate use of TEMP than is the case when such selections are made on the basis of unvalidated interviews conducted by untrained managers.

#### VIII. SUMMARY AND CONCLUSIONS

Under a grant from the U. S. Public Health Institute, project personnel constructed a manual for the quasi-objective evaluation of ten TAT protocols to be used primarily with normal subjects in the industrial setting. The end product of the initial research was a manual entitled "Thematic Evaluation of Managerial Potential."

In a second study, jointly sponsored by the Office of Naval Research and the Personnel Research and Development Corporation, 50 evaluators volunteered to evaluate ten protocols each. Five of these were common to all 50 raters. The other five protocols rated by each of the 50 raters were selected from a pool of 125 sets of protocols in such a way that each set of protocols was rated by two of the evaluators. The reliability of the 18 scales was measured by intra-class correlations. Examination of the scatter plots revealed that certain of the scales could be improved by revision. Ten of the 18 scales met the

arbitrary criterion of an intra-class correlation between raters of .30. After revision, an 11th also met this criterion. Seven of the 18 scales were discarded as not being sufficiently reliable. The other 11 were retained in a revised manual. Of these, the scale steps in three were changed to some degree as a result of the examination of the scatter plots of the two sets of evaluations.

The revised manual is being prepared for publication in book form for release as an experimental selection device whose validity remains to be established, both for particular jobs and for specific evaluators.

## APPENDIX A

Rating Scales Employed in Original TEMP Manual

## 1.0 Accuracy of Perception

A. Rarely sees objects or people in same way as others. Usually distorts their general features or characteristics. Very unlikely to agree with or communicate well with others.	B. Record shows some inaccuracies in perception that are of major importance.	C. Frequent inaccuracies in perception that are of a minor nature. But taken as a whole, these make serious lapses.	D. May show some minor inaccuracies, but usually sees things the way others do.	E. Sees objects or features in unusual, but realistic, acceptable manner. Original and believable.
--	---	---	---	--

## 2.0 Amount of Productivity

A. Does not meet task requirements. Does not produce in satisfactory quantity. Is apathetic, seems unaware of, or resistant toward, work standards.	B. Scanty, sparse. Meets minimal requirements but below normal in total output. Does not take pains to fulfill standards.	C. Produces usual amount. Meets ordinary standards. Is not impressive in meeting standards, but does so.	D. Produces in rather large quantity. Exceeds ordinary productivity. Gives generously of energies to tasks at hand.	E. Over-produces. Goes beyond the bounds of the task. Perseverates, over-elaborates in tasks. Tries too hard, over-fulfills.
---	---	--	---	--

### 3.0 Consistency of Productivity

A. Output decreases with time. "Runs down" fast. Accumulates task-resistance.	B. Output generally even and steady. Can be counted on to apply steady effort.	C. Output increases with time. Becomes more effective and productive as he proceeds.	D. Fluctuates rather widely. Shows significantly large variations. Erratic and unreliable. Can produce effectively on some occasions but falls into slumps. Can be quite inactive at times.	

### 4.0 Whole versus Detail Approach to Problems

A. Concentrates heavily on one or two details. Does not integrate details. Misses implications of problems.	B. More concerned with details than broad aspects, but some integration of details is shown.	C. Addresses himself to the broad aspects of a problem. Supports the overall view using all the major, important details.	D. Emphasizes a general view, but overlooks some important details.	E. Neglects details. Focuses only on broad aspects. Over-simplifies. Glosses over details.

## 5.0 Originality of Expression

A. Unoriginal, stereotyped ideas and manner of treatment. Very ordinary.	B. Ideas are basically conventional, but occasionally expressed in a somewhat unusual way.	C. Has some unusual ideas of good quality but not deep nor richly elaborated.	D. Definitely original in a constructive way. Ingenious ideas and style of expression.	E. Very imaginative but not in a constructive way. Originality based on distortion or individual preoccupation.	F. Peculiar ideas, out of touch with reality. May be overly abstract or symbolic. Strange thinking.

## 6.0 Emotional Control

A. Tends to be stiff and "cold" in outlook or approach. Over-controlled in emotions.	B. Tends to be rather passive and bland in approach. Seldom is "moved" to show any but superficial, pleasant emotions.	C. Appropriately emotional. Is either lightly or deeply responsive to people and events as is called for.	D. Tends to respond openly to emotional situations. Shows some inclination to intense feeling but strength of reaction is not extreme.	E. Inclined to erratic reactions, highly volatile emotions. Easily moved to powerful reaction, openly displayed.	

## 7.0 Involvement in Work

<p>A. Little involvement in work. Does as little as he can get away with. Works only to meet minimum demands of job and life.</p>	<p>B. Works more from a sense of duty or obligation than from a real sense of challenge. May be highly involved at times but the work is not directly satisfying. Motivated by conscience or fear of the consequences of not performing, rather than from any real liking for work.</p>	<p>C. Moderate involvement in work. Derives some pleasure from work itself but extrinsic rewards such as income, social life are more important. Can work hard especially when rewards are greater but doesn't become overly absorbed in the work, per se.</p>	<p>D. Highly involved in work. Derives much direct, genuine satisfaction from work itself. Enjoys the challenge, takes his work seriously and takes pride in accomplishing worthwhile results.</p>	<p>E. Excessive involvement in work. So stimulated and challenged by his work that other needs and life demands are neglected. Willing to make great sacrifices in other areas for sake of work goals. Doesn't know when to "put down" a problem and relax or turn his efforts to other matters.</p>
---	---	--	--	--

## 8.0 Need for Clear Assignment and Direction

<p>A. Depends strongly on instruction from others. Needs to have duties and actions spelled out in detail. Cannot work on the basis of general instruction.</p>	<p>B. Works best in a well-defined situation. When goals and policies are clearly laid out, he can fill in with routine procedures. Has trouble if only given general instruction.</p>	<p>C. Carries out normal job activities with slight need for specific instruction. Can work acceptably well from general assignments.</p>	<p>D. Likes to work things out for himself. Works best when allowed to work on general plans as well as specific procedures. Needs little instruction.</p>	<p>E. Has trouble sticking with plans or procedures whether general or specific. Distractible. Changes direction or attention rapidly, according to what comes along.</p>
---	--	---	--	---

## 9.0 Need for Emotional Support in Work

Needs constant reassurance and support from others, even in routine activities.	B. Needs much reassurance and support in new undertakings. Is independent in routine work.	C. Needs some support and consultation in new ventures of a complex nature. Needs no help in ordinary matters.	D. Needs only minimum assistance. Tackles complex tasks by himself. Does his own planning. Figures out his own solutions, but readily makes use of help when appropriate.	E. Rash, overconfident, desires no help. Dislikes consulting with others. Acts as if he needs no support.
---	--	--	---	---

## 10.0 Reaction to Failure and Frustration

A. Becomes easily discouraged and depressed in the face of obstacles. Has low tolerance for frustration and is dominated by it. Tends to accept the frustration as insurmountable and may reduce his level of aspiration because of a feeling of failure. Frequently blames himself for his difficulty.	B. Becomes highly impatient or angered, tending to blame others or external objects for the thwarting. May avoid assuming responsibility for overcoming the obstacle and attempt to escape possible consequences of failure.	C. Reacts with normal disappointment to frustration, but makes determined attempt to remove the blocking object. May need a little encouragement and support but generally can adjust well and tackle the problem intelligently.	D. Can withstand a good deal of frustration and takes it in stride. Responds by determination to overcome the obstacle and achieve the goal. Is even challenged by the problem presented and heightens his efforts to remove the barrier.	E. Is overly optimistic about his ability to resolve the situation or denies the existence of any frustration. May fail to see seriousness of the situation or, can put in inappropriate effort and not know when to give up and accept his failure. May be blind to realities or "overly tolerant" of frustration.
---	--	--	---	---



## 11.0 Level of Ambition

A. Contented with present level and/or earnings; doesn't care about advancement in his work; has found his niche in business.	B. Wants only a little more than he now has; perhaps a level higher.	C. Wants to advance a few levels but doesn't aspire to a high position.	D. Wants a high level but not necessarily top position; would like substantially more income or status.	E. Aspires to top position in status or earnings; wants a great deal more than he has.

## 12.0 Intensity of Ambition

A. Contented with present level and/or earnings. Doesn't care about advancement. Has found his niche in business.	B. Has a higher goal in business but is unable or unwilling to do more than dream about it. Puts forth no real effort to advance.	C. Efforts to advance are moderate or sporadic. Unwilling to put forth any consistent effort or make much sacrifice to achieve the level to which he aspires.	D. Efforts to achieve advancement are strong and consistent. Shows willingness to invest time and resources in getting ahead.	E. Intensely determined to reach his goal. Weighs nearly everything he does in terms of his advancement. Willing to make sacrifices and/or enemies to get ahead.

### 13.0 Relations with Superiors

<p>A. Resents superiors. Likely to be argumentative and stubborn. Hard for a superior to get along with.</p>	<p>B. Somewhat difficult for his superiors. Shows some tendencies to rebellion or stubbornness.</p>	<p>C. Able to relate positively to superiors. Takes direction without resentment or competitive strivings.</p>	<p>D. Inwardly resentful of superiors, but is submissive and deferent in behavior.</p>	<p>E. Passively accepts superior's authority. Deferent and submissive in behavior.</p>
--	---	--	--	--

### 14.0 Ability to Organize and Direct Others

<p>A. Poor in planning or organizing. Unable to delegate or assign work to others in a systematic way or to coordinate the efforts of the group. Is weak in setting-up and working toward goals, schedules, and deadlines. Laissez-faire approach; poor controls or authority exercised. Lacks desire to plan.</p>	<p>B. Inconsistent or erratic in organizing or coordinating work of his subordinates. Plans and goals often only vaguely formulated and not always followed-up. Vacillates between expecting subordinates to do their work with minimum direction and being very firm and direct when he does decide what is to be done. Assignments given out impulsively.</p>	<p>C. Plans and coordinates activities and work of others effectively. Can schedule, delegate and assign activities smoothly, setting reasonably high standards and a well-balanced work pace. Able to supply controls and directions when needed; good perspective on total task of function of group.</p>	<p>D. Plans and coordinates activities of his group efficiently, but tends to be inflexible and strictly bound by rules and regulations. Often interferes with work he has assigned by supervising too closely.</p>	<p>E. Maintains excessive control and authority. Delegates very little to subordinates and relinquishes little or no authority to them. May be a well-organized planner, but total direction of his work group is one-sided because he lacks trust in subordinates or prefers to do things personally. "I am the boss" type.</p>
--	---	---	---	--

## 15.0 Consideration for Subordinates

- |  |  |  |  |  |
|--|--|--|--|--|
| <p>A. Emphasizes status differences and lacks warmth in his dealings with subordinates. Treats subordinates in a superior, condescending, even demeaning way. Likely to "downgrade" their ideas and to treat their attitudes and feelings as being relatively unimportant and unworthy of importance. Little or no particular participation permitted.</p> | <p>B. Generally avoids thinking of subordinates and their problems. Expects them to do their jobs with little attention from him. Tends to be a "lone wolf" who is disinclined to carry responsibility for the efforts and problems of others. Is cool and impersonal with subordinates.</p> | <p>C. Inconsistent. Shows concern for subordinates ideas and feelings on some occasions, but disregards such consideration at other times. Tends to be overly involved in meeting production goals and insensitive to needs of subordinates.</p> | <p>D. Usually seeks out or takes account of the ideas and feelings of subordinates in making decisions. Encourages and supports their efforts. Shows genuine interest in their development and in their personal needs, but balances this approach against performance requirements.</p> | <p>E. Excessively concerned with feelings of his workers and with harmony in the work group. Concerned about possible disagreements or lack of approval to point where it interferes with his effectiveness.</p> |
|--|--|--|--|--|

## 16.0 Insight into Others

- |  |  |  |  |   |
|--|--|--|--|---|
| <p>A. Avoids thinking about others in any personalized way. Ignores issues about feelings and motives.</p> | <p>B. Tends to stereotype others; oversimplifies; classifies with "good - bad" labels. Tends to be intolerant.</p> | <p>C. Has ability to understand ordinary feelings and motives; but only moderately perceptive.</p> | <p>D. Consistently able to understand other people well. Takes time to evaluate people as individuals. Perceptive and sensitive to differences among them.</p> | <p>E. Dwells too much on analysis of others. Spends too much time evaluating them. Preoccupied with feelings and motives.</p> |
|--|--|--|--|---|

## 17.0 Social Skills

A. Crude and rough. Lacks even ordinary graces. Offensive.	B. Definitely awkward and ill-at-ease. Has trouble in many social situations. Inoffensive.	C. Shows moderate amount of poise and skill. Can adapt to various situations fairly well. Pleasant.	D. Shows definitely good poise and skill. Can adapt to diverse groups and situations with considerable ease. Natural and likeable.	E. Highly polished and smooth. Poised and graceful in nearly any situation. Natural and charming.	F. Tends to be pretentious and snobbish. Tries too hard to impress others. Offensive.

## 18.0 Social Dominance

A. Socially unassertive. Seldom shows any initiative or dominance.	B. Shows consistent self assertion. Can often take the initiative in various situations.	C. Shows frequent initiative and self-assertiveness but is not objectionable in doing so. Contributes by his influence and dominance.	D. Too dominating. Attempts to control or influence nearly any situation. Objectionably self-assertive.		

APPENDIX B

PERSONNEL RESEARCH AND DEVELOPMENT CORPORATION

USPHS Grant #M-2158 The TAT in Predicting Normal Adjustment

EXPERIMENTAL ANALYSIS PROTOCOL

Rater Name \_\_\_\_\_ Rater No. \_\_\_\_\_ Date No. \_\_\_\_\_

Supplementary Protocol Information

Age in years \_\_\_\_\_

Time \_\_\_\_\_

Education:

Marital Status:

Less than High School \_\_\_\_\_

Single \_\_\_\_\_

Some High School \_\_\_\_\_

Married \_\_\_\_\_

High School Graduate \_\_\_\_\_

Divorced \_\_\_\_\_

Some College \_\_\_\_\_

Separated \_\_\_\_\_

College Graduate \_\_\_\_\_

Remarried \_\_\_\_\_

Post Graduate \_\_\_\_\_

Widowed \_\_\_\_\_

Explanatory Note

This protocol has been faithfully typed from original handwritten copy. You will discover many instances of apparent typographical error, spelling, punctuation, omissions, capitalization and other conventions of form. However, these are intentional reproductions of the subject's original writing. Where the word "JOHN DOE" is found at the beginning or end of a story, this indicates that the subject has signed his name to the story. Some errors and changes made by the subject in writing out his story have been typed in parenthesis with the following code used to indicate the nature of the error or change.

(+) = Additions or insertions

(E) = Readable Erasures

(#) = Readable Crossouts

(X) = Obliterated Erasures

(-) = Undecipherable Crossouts

"JOHN DOE" = Subject has signed his name

Example:

Now he (E is) has gone to (# the) country.

## APPENDIX C

### THEMATIC EVALUATION of MANAGERIAL POTENTIAL

(National Institute of Mental Health  
Grant M-2158)

#### Instructions for Completing Evaluations

On the following pages are eighteen rating scales with which you are to evaluate ten cases, or the protocols, contained in this envelope. Each scale has been divided into four, five, or six divisions with a paragraph describing each. You are to rate each protocol on all eighteen scales before proceeding to the next protocol.

Before making your ratings, read Chapter IV of the manual for directions on how to use the manual for making ratings. Make sure to read the entire manual once before beginning to make your ratings.

In rating the protocols, proceed as follows: First, read each of the scale level descriptions. Then decide which category best fits your judgment of the protocol, as determined by your study of indicators present in the story. Enter the three-digit case number in the space above the description you have chosen. Always put the case number in one of the spaces, never on the line between two categories. Be sure to rate each protocol on all eighteen scales. The value of your ratings to the research project depends on your following the instructions exactly.

As you complete the work on a particular protocol, and have made your ratings, fill in the space on the envelope following the case number with the date and your initials.

Illustration of how ratings should be noted:

#### 3.0 CONSISTENCY of PRODUCTIVITY

105		017 099	
Output decreases with time.	Output generally even and steady.	Output increases with time.	Fluctuates rather widely.

In rating Case No. 105, the evaluator decided that the protocol indicated that this person's productivity would fall off with time. Therefore, he entered the case number, 105, in the space over this category. Cases 017 and 099 fell into the same category, so both numbers were entered in the appropriate space.

When all protocols have been completed, each scale should have ten three-digit numbers entered. Before returning the materials to PRADCO, please check to see that no omissions have been made. Then place the rating scales back in the white envelope they were contained in and return this envelope in the manila, self-addressed return envelope provided for this purpose. You are to keep the manual and the case records.

## REFERENCES

1. Anderson, H. H., & Anderson, Gladys L. An introduction to projective techniques. New York: Prentice-Hall, 1951.
2. Aron, B. A manual for analysis of the Thematic Apperception Test. Berkeley, California: Willis E. Berg, 1949.
3. Bellak, L. A guide to the interpretation of the Thematic Apperception Test. New York: Psychological Corp., 1951.
4. Combs, A. W. A comparative study of motivations as revealed in Thematic Apperception stories and autobiography. J. clin. Psychol., 1947, 3, 65-75.
5. Combs, A. W. The validity and reliability of interpretation from autobiography and the Thematic Apperception Test. J. clin. Psychol., 1946, 2, 240-247.
6. Dana, R. H. An application of objective TAT scoring. J. proj. Tech., 1956, 20, 159-163.
7. Eron, L. D. A normative study of the Thematic Apperception Test. Psychol. Monogr., 1950, 64, No. 9, (Whole No. 315).
8. Friedman, I. Objectifying the subjective -- A methodological approach to the TAT. J. proj. Tech., 1957, 21, 243-247.
9. Gardner, B. What makes successful and unsuccessful executives? Advanc. Mgmt., 1948, 13, 116-125.
10. Gaudet, F. J. (with A. R. Carli) Why executives fail. Personnel Psychol., 1957, 10, 7-21.
11. Goodman, M. An indirect validation of a TAT scoring manual. J. clin. Psychol., 1952, 8, 149-154.
12. Harrison, R. Studies in the use and validity of the Thematic Apperception Test with mentally disordered patients. II. A quantitative validity study. III. Validation by the method of "blind analysis." Charact. & Pers., 1940, 9, 122-138.
13. Hartman, A. A. An experimental examination of the Thematic Apperception Test in clinical diagnosis. Psychol. Monogr., 1949, 303, (Whole No. 303).
14. Henry, W. E. Executive personality and job success. Personnel Series, 1948, 120, 3-13.
15. Henry, W. E., & Gardner, B. B. Personality evaluation in the selection of executive personnel. Pub. Personnel Rev., 1949, 10, 67-71.
16. Henry, W. E. The business executive: the psychodynamics of a social role. J. Pers., 1951, 19, 312-323.

17. Henry, W. E. The analysis of fantasy. New York: John Wiley & Sons, 1956.
18. Holtzman, W. H., & Sells, S. B. Prediction of flying success by clinical analysis of test protocols. J. abnorm. soc. Psychol., 1954, 49, 485-490.
19. Horst, Paul. A generalized expression for the reliability of measures. Psychometrika, 1949, Vol. 14, No. 1, 21-31.
20. Lindzey, G. TAT: assumptions and related empirical evidence. Psychol. Bull., 1952, 49, 1-25.
21. McClelland, D. C., Atkinson, J. W., Clark, R. A., & Lowell, E. L. The achievement motive. New York: Appleton-Century-Crofts, 1953.
22. Morgan, C. D. & Murray, H. A. A method for investigating fantasies: The Thematic Apperception Test. Arch.Neurol.Psychiat., 1935, 34, 289-306.
23. Murray, H. A. Explorations in personality. New York: Oxford Univ. Press, 1938.
24. Murray, H. A. Thematic Apperception Test. Pictures and manual (20 pp.) Cambridge, Mass.: Harvard Univ. Press, 1943.
25. Nevis, E. C., Wallen, R. W., & Stickel, E. G. The Thematic Evaluation of Managerial Potential. Cleveland, Ohio, 1959. Personnel Research and Development Corporation (Out-of-print).
26. Ricciuti, H. N., & Schultz, D. G. Level of aspiration measures and self-estimates of personality in relation to achievement motivation. Technical report, Office of Naval Research Contract No. Nonr 694(00). Princeton, N.J.: Educational Testing Services, 1958.
27. Sadacca, R., Clark, R. A., & Ricciuti, H. N. Content analysis of achievement motivation protocols: a working manual. Technical report, Office of Naval Research Contract No. Nonr 694(00). Princeton, N.J.: Educational Testing Service, 1957.
28. Stein, M. I. The Thematic Apperception Test: an introductory manual for its clinical use with adult males. Cambridge, Mass.: Addison-Wesley, 2nd Ed., 1954.
29. Stern, G. G., Stein, M. I., & Bloom, B. S. Methods in personality assessment. Glencoe, Ill.: The Free Press, 1956.
30. Thorndike, Robert L. Personnel Selection. New York: John Wiley & Sons, 1949.
31. Tomkins, S. S. (with collaboration of Elizabeth J. Tomkins). The Thematic Apperception Test: the theory and technique of interpretation. New York: Grune & Stratton, 1947.
32. Ullman, L. P., & McFarland, R. L. Productivity as a variable in the TAT protocols - a methodological study. J. proj. Tech., 1957, 21, 80-87.